Abstract

The screening and testing program the US Environmental Protection Agency (EPA) is currently developing to detect endocrine-disrupting chemicals (EDCs) is described. EDCs have been shown to alter the following activities: hypothalamic-pituitary-gonadal (HPG) function; estrogen, androgen, and thyroid hormone synthesis; and androgen and estrogen receptor-mediated effects in mammals and other animals. The value and limitations of mammalian in vivo assays are described that involve the use of the laboratory rat, the EPA Endocrine Disruptor Screening and Testing Advisory Committee species of choice. The discussion includes the evaluation of high-priority chemicals positive in the Tier 1 Screening (T1S) battery, and of subsequent testing in the Tier 2 (T2) battery, with additional short-term screening assays proposed for use in T1.5 to eliminate any uncertainty about T1S results. Descriptions include in vivo uterotrophic assay, which detects estrogens and androgens; and the pubertal female assay, which assesses steroidogenesis, antithyroid activity, antiestrogenicity, and HPG function; and the Hershberger assay, which detects the weight of androgen-dependent tissues in castrate-immature-male rats (antiandrogens). Of the several alternative mammalian in vivo assays proposed, a short-term pubertal male rat assay appears most promising for inclusion in T1 or T1.5. An additional in utero-lactational screening protocol is being evaluated, but appears to be better suited for T1.5 or T2 due to the size, complexity, and duration of the assay. The adult intact male assay, also proposed as an alternative for T1, attempts to identify EDCs in a hormonal battery, but has limited value as a screen due to lack of sensitivity and specificity. For Tier 2 testing, the number of endocrine-sensitive endpoints and offspring (F1) examined in multi-generational tests must be thoughtfully expanded for EDCs on a mode-of-action-specific basis, with consideration given to tailoring T2 based on the results of T1S.

Key Words: androgens; antiandrogens; endocrine disruption; estrogens

Introduction

The potential effects of endocrine-disrupting chemicals (EDCs) on human health and the causal effects of EDCs demonstrated in wildlife constitute a major concern among the public and the scientific community. In 1994, due to potential gaps in the current testing of chemicals, the US Environmental Protection Agency (EPA) received a mandate under the Food Quality Protection and Safe Drinking Water Acts to develop a screening and testing program for endocrine effects. Some of the impetus for the mandate arose from a Work Session in 1991 on Chemically Induced Alterations in Sexual Development: The Wildlife/Human Connection (Colborn and Clement 1992), which stated, “Many compounds introduced into the environment by human activity are capable of disrupting the endocrine system of animals, including fish, wildlife, and humans. Endocrine disruption can be profound because of the crucial role hormones play in controlling development” (Colborn 1994; Colborn and Clement 1992, p. 1). The scientific consensus of the workshop “estimated with confidence” that impairments in humans have resulted from exposure to en-

Abbreviations used in this article: ADME, absorption, distribution, metabolism, and excretion; AR, androgen receptor; EDC, endocrine-disrupting chemical; EDSP, Endocrine Disruptor Screening Program; EDSTAC, Endocrine Disruptor Screening and Testing Advisory Committee; EPA, Environmental Protection Agency; ER, estrogen receptor; HPG, hypothalamic-pituitary-gonadal; HTPS, high-throughput prescreening; ICCVAM, Interagency Coordinating Committee on the Validation of Alternative Methods; insl-3, insulin-like 3 peptide; OECD, Organisation for Economic Cooperation and Development; OPPTS, Office of Prevention, Pesticides and Toxic Substances; QSAR, quantitative structure activity relationship.

The research described in this article has been reviewed by the National Health Environmental Effects Research Laboratory, US Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the Agency nor does mention of trade names or commercial products constitute endorsement or recommendation for use.
Endocrine disruptors. Laboratory studies corroborate the abnormalities of reproductive development observed in the field and, in some cases, define toxic mechanisms causing adverse effects. Among the chemicals of concern are pesticides and industrial chemicals, pharmaceuticals, phytochemicals, "inert" ingredients, food supplements, personal care products, and "natural" products, often referred to as "nutraceuticals."

In response to the 1996 legislative mandate for an endocrine screening and testing program, the EPA formed the Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC), which proposed a tiered screening and testing strategy for EDCs in its final report in 1998 (http://www.epa.gov/scipoly/oscpendo/history/finalrpt.htm). The EDSTAC proposal included the following: (1) a process to prioritize chemicals for evaluation and recommendations, for (2) screening (Tier 1), and for (3) testing (Tier 2) batteries. The chemical “universe” under consideration includes more than 80,000 chemicals, of which only a subset of high-priority chemicals would first enter the screening program. Initially, EPA planned to estimate the ability of the chemicals to interact with steroid hormone receptors in the prioritization process, and to examine using either the quantitative structure activity relationship (QSAR), for chemicals that bind steroid receptors, or high-throughput prescreening (HTPS), using hormone-dependent gene expression assays. However, neither robust validated HTPS methods nor QSAR models for steroid receptors interactions have yet been developed. These methods are currently unavailable for the prioritization of chemicals for screening.

The screening battery recommended by EDSTAC was designed to detect the following: alterations of hypothalamic-pituitary-gonadal (HPG) function; estrogen, androgen, and thyroid hormone synthesis; and androgen receptor (AR)- and estrogen receptor (ER)-mediated effects in mammals and other taxa. Based on a “weight-of-evidence” analysis, chemicals positive in Tier 1 would be considered as potential EDCs and subjected to testing (Tier 2). Equivalent effects in Tier 1 could be replicated or evaluated further in additional short-term assays before more extensive Tier 2 testing was initiated. Tier 1 should include assays sensitive enough to detect EDCs, whereas issues of “dose-response, relevance of the route of exposure, sensitive life stages and adversity” would be resolved in the Tier 2 testing phase.

Based largely on the EDSTAC recommendations, the EPA Office of Prevention, Pesticides and Toxic Substances (OPPTS) designed the Endocrine Disruptor Screening Program (EDSP), and the Agency began to implement elements of the program in 1998 (the report to Congress is available at http://www.epa.gov/scipoly/oscpendo/index.htm#currentstatus). Meeting summaries, background documents, final reports of studies performed on contract for EPA, and slides that summarize the results of the studies are available at http://www.epa.gov/scipoly/oscpendo/assayvalidation/meetings.htm. Many current efforts related to the uterotrophic and Hershberger assays are being coordinated internationally with the Organisation for Economic Cooperation and Development (OECD) (described at http://www.oecd.org/ehs/ENDOCRIN.HTM). OECD is also evaluating in vitro assays for endocrine screening.

Due to limitations of the in vitro assays, it is necessary to include in vivo assays in the screening battery. In vitro assays alone cannot account for absorption, distribution, metabolism, and excretion (ADME), and they yield false-negative and false-positive results. In particular, many false-negatives result from the inability of the in vitro assays to activate EDCs metabolically. In addition, in vitro assays at high concentrations lack specificity and produce many false-positive results. Because in vivo assays not only account for ADME but also can integrate all of the endocrine and nonendocrine toxicities, the combination of both in vivo and in vitro assays is recommended for screening.

In vivo assay development and validation are supported by recommendations from the Endocrine Disruptor Methods Validation Subcommittee, as described on the EDSP website; whereas in vitro assay development and validation efforts are reviewed by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) of the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (http://iccvam.niehs.nih.gov/methods/endocrine.htm). At the request of the EPA, ICCVAM convened an independent panel of scientists in 2002 to assess the validation status of in vitro assays proposed as screens to identify EDCs that acted via the ARs or ERs, including binding and transcriptional activation assays. This panel proposed minimum performance standards for in vitro assays and provided lists of chemicals that could be used to validate the assays (see Final ICCVAM Evaluation Report at http://iccvam.niehs.nih.gov/methods/endodocs/edfinrpt.htm).

EDSTAC recommended the laboratory rat as the species of choice for the endocrine screening and testing assays. In this article, we discuss the use of the laboratory rat as an animal model to detect endocrine alterations of HPG function, estrogen and androgen hormone synthesis, and AR- and ER-mediated effects. The reader is referred to a recent workshop review of the assays of thyroid function, which provides an excellent overview of this pathway (Jahnke et al. 2004).

**Rat and Human Reproductive Function: Similarities and Differences**

Research on the reproductive physiology and endocrinology of the rat as an experimental animal began in the 1930s. Since then, the species has been more thoroughly characterized in these research fields than any other laboratory animal model, and it has been the species of choice for multigenerational testing studies for several decades. In 1962, an expert panel published several standardized and validated short-term assays for the purpose of screening for...
endocrine activity in the rat (Dorfman 1962). Among those screening assays were the uterotrophic assay, designed to detect estrogens, and the Hershberger assay (Hershberger et al. 1953), designed to detect androgens and antiandrogens. The pubertal male (Gray et al. 1988a, 2001; Monosson et al. 1999; Stoker et al. 2000) and the female rat assays (Goldman et al. 2000; Gray et al. 1988b, 1999a, 2001; McIntyre et al. 2000) were developed at the EPA in the 1980s; and the adult intact male rat assay (Cook et al. 1993) was first proposed in 1993. In 1996, a workshop sponsored jointly by the EPA, the Chemical Manufacturers Association, and the World Wildlife Fund recommended the use of both in vitro and in vivo assays to screen chemicals for endocrine activity (Gray et al. 1997a). The approach proposed by this panel was very similar to the proposed EDSTAC Tier I Screening battery.

The utility of the laboratory rat as an animal model for endocrine screening is evident in the literature, which not only contains a wealth of information about the endocrine control of reproductive function and development, but also is rich with evidence on the effects of endocrine-disrupting toxicants on this system. Short-term in vivo assays were first used to screen pesticides for estrogenic activity in the 1960s. To extrapolate effects in the rat to humans and other vertebrate species, we must use the wealth of information that is available in the literature, some of which dates back to the 1930s. We must recognize not only the similarities (Table 1a) among mammals in reproductive endocrine function, but also the differences (Table 1b).

### Cellular and Molecular Levels

At the cellular and molecular levels, reproductive endocrine function is highly conserved among mammalian species. All mammals appear to have a single AR and two ERs—alpha and beta. In contrast, some fish have two ARs, and at least four ERs have been described in some species. In all vertebrates, the steroid hormone receptors act as nuclear transcription factors that regulate hormone-sensitive gene expression. The AR and ER bind natural and synthetic ligands, recruit coactivators, and form homodimers. This complex is translocated to the DNA, where they bind to hormone response elements on the genes, resulting in alterations of gene expression, protein synthesis, cell growth, and function. These receptors show a high degree of sequence homology among mammals, especially in the ligand binding domain. It is therefore reasonable to expect that a chemical that acts as an antiestrogen or antiandrogen in an assay using the laboratory rat would also produce predictable reproductive alterations in humans. Species differences in responses to man-made chemicals would be the rare exception, rather than the rule, at the level of androgen or estrogen receptor binding and transactivation.

The steroidogenic pathways involved in synthesis of gonadal and adrenal hormones are also highly conserved, although different species have “preferences” for specific steroidogenic pathways. Drugs that inhibit cytochrome P450 enzymes in these steroid synthesis pathways can be expected to affect steroid hormone synthesis in both rats and humans. Generally, these chemicals also are effective in fish, amphibians, birds, reptiles, and some invertebrates. For example, drugs that inhibit aromatase and other P450 enzymes in the rat produce antiestrogenic effects in fish, and inhibit ecdysone synthesis and reproductive function in invertebrates (Mu and LeBlanc 2002). However, it is important to note that the $K_c$ values of P450 inhibitors can show considerable variation between mammals and fungi, and such species differences may also exist among vertebrates.

### Anatomy and Physiology

At levels of biological organization higher than the cellular and molecular level, the anatomy and physiology of the endocrine system also are highly conserved among vertebrates. However, as the level of biological complexity increases, species differences also increase. For some hormones, like estradiol and testosterone, the functions are

<table>
<thead>
<tr>
<th>Table 1a Examples of reproductive physiology similarities among humans and rats</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Steroid hormone control of reproductive function relies on testosterone, dihydrotestosterone, estradiol, and progesterone.</td>
</tr>
<tr>
<td>• CNS-hypothalamic secretion of GnRH controls pituitary release and synthesis of FSH and LH.</td>
</tr>
<tr>
<td>• FSH and LH regulate germ cell development after puberty, LH surges induce spontaneous ovulation in the female, LH regulates testis Leydig cell testosterone production.</td>
</tr>
<tr>
<td>• Placental support of embryos. Placenta and fetal unit also produce hormones critical for pregnancy maintenance after the first week.</td>
</tr>
<tr>
<td>• Hormonal regulation of uterine function and onset of delivery.</td>
</tr>
<tr>
<td>• Androgens required to maintain male spermatogenesis and secondary sex characteristics.</td>
</tr>
<tr>
<td>• Hormone-dependent mating and other sexually dimorphic behavior. “Rough and tumble” play behavior is sexually dimorphic behavior being imprinted by early androgens.</td>
</tr>
<tr>
<td>• Lactation under complex hormonal regulation.</td>
</tr>
<tr>
<td>• Dramatic endocrine changes resulting from CNS-HPG maturation responsible for puberty in males and females. Females generally attain puberty at an earlier age than males of the same species.</td>
</tr>
</tbody>
</table>

---

$^a$CNS, central nervous system; GnRH, gonatropin-releasing hormone; FSH, follicle-stimulating hormone; LH, luteinizing hormone; HPG, hypothalamic-pituitary-gonadal.
highly conserved among vertebrates and are important for reproduction. However, the role of prolactin is extremely diverse among vertebrates.

Anatomically, both rat and human reproductive function involves central nervous system hypothalamic-pituitary control of gonadal function. All eutherian mammals share common endocrine-modulated traits such as internal fertilization and intrauterine development. Hormones are required to induce implantation, maintain pregnancy, initiate parturition, and induce lactation and maternal and reproductive behaviors. However, as the rat and human have very different reproductive strategies, significant differences exist in specific endocrine-regulated events. Selected examples of these are shown in Table 1b.

### In Vivo Mammalian Assays

In vivo assays are required in TIS due to the following major limitations of in vitro assays: (1) In vitro assays cannot account for metabolic activation of xenobiotics resulting in “false-negative” responses (positive in vivo, but negative in vitro). (2) These assays also can yield false-positive responses when chemicals are examined at in vitro concentrations orders of magnitude that exceed those seen in vivo. (3) In vitro assays can produce false-positive responses at high concentrations due to a lack of specificity as assay conditions deteriorate. Although such false-positives can be eliminated by experimental measurement of Kᵢ values, few in vitro screening strategies include Kᵢ determinations to ensure that the effects seen in binding and gene expression assays are the result of competitive inhibition of ER or AR. As a result, a high percentage of screened chemicals that are determined to be “positives” in vitro may not be true receptor ligands. For this reason, QSAR models of ER or AR binding based on IC₅₀ values or relative binding affinities rather than Kᵢ values may have a considerable amount of misinformation in the training set data used to program the computer to recognize ligands. (4) Finally, because in vitro screening assays are unable to integrate the endocrine re-

---

**Table 1b Examples in which the reproductive strategy of the rat differs from that of the human**

1. The rat is a short (22.5-day) gestation species. Pregnancy in humans is 9 mo.
2. The rat placenta lacks aromatase; estrogen is produced during pregnancy by the ovary. Human placental tissue expresses high levels of aromatase.
3. In the rat, sexual differentiation of the reproductive tract is perinatal, whereas central nervous system (CNS) sexual differentiation is a postnatal event, regulated to a great degree by aromatization of testosterone to estradiol (play behavior, an exception, is androgen dependent in both rats [Hotchkiss et al. 2002, 2003] and humans [Hines 2003]).
4. The rat has a 4- to 5-day estrous cycle, with no functional corpus luteum. The estrous cycle can be monitored easily by examining daily cytology. The female rat displays sexual receptivity only during estrus after “lights out” after a proestrus vaginal smear. This behavior is exquisitely dependent on estrogen followed by progestosterone. Humans have a menstrual cycle approximately 28 days in duration and do not display periods of peak behavioral estrus during the cycle. Corpora luteal function is sustained for approximately 10 days by mating-induced cervical stimulatory prolactin surges in rats, whereas the human menstrual cycle has a spontaneous luteal phase of 10 to 14 days after ovulation.
5. Male rat sex behavior can be induced by estrogens and involves multiple series of ejaculations in a single mating. Mating involves approximately 10 mounts, with intromission before each ejaculation, followed by a postejaculatory interval before the onset of the next series. In nonhuman primates and presumably humans, male sex behavior is androgen mediated.
6. Both ovaries spontaneously release several ova in response to a luteinizing hormone surge into separate uterine horns, each with a separate cervix in the rat; whereas in women, a single ovum is typically ovulated during each cycle.
7. Pregnancy is easily disrupted by estrogens in rats, but not in humans. Rats, unlike humans, are a litter-bearing species. Most strains used for toxicology testing have litters of 10 to 12 pups. Spontaneous reproductive malformations are very rare in the rat, whereas in humans, some malformations such as cryptorchidism occur in 3% of newborn boys.
8. Spermatogenesis begins at approximately 5 days of age in the rat; the spermatogenic cycle is about 53 days of age, and sperm appear in the epididymis at about 55 days of age. In humans, spermatogenesis begins during puberty at 10 to 14 yr of age, and the entire spermatogenic cycle is approximately 75 days in duration.
9. Puberty in the rat (as measured by the age at vaginal opening and the onset of estrous cyclicity) occurs at about 32 days of age in females and 42 days of age (as measured by preputial separation an androgen-dependent event) in male SD and LE rat strains. In humans, puberty occurs at 9 to 12 yr of age in girls, and 10 to 14 yr of age in boys.
10. Fertility begins to decline in the female rat at about 6 mo of age, especially if never mated and allowed to cycle continuously. Fertility begins to decline in women at about 35 yr of age, and at 40 yr of age, approximately 50% of women are infertile.
sponses seen in the whole organism, the relationships between endocrine toxicity and other systemic effects cannot be simulated in vitro.

To avoid the limitations described above, the EDSTAC proposed three short-term in vivo mammalian assays for the Tier 1 Screening Battery: the uterotropic, Hershberger, and pubertal female rat assays (Gray 1998b; Gray et al. 1997a).

**Uterotropic Assays**

In the EDSTAC T1S, estrogen agonists and antagonists are detected in a 3-day uterotropic assay using subcutaneous administration of the test compound. Based on the evaluation of four variations of the uterotropic assay protocol in the OECD interlaboratory studies, all of the protocols have produced acceptable responses without regard to rat strain, diet, or housing conditions (Kanno et al. 1001, 2002, 2003a,b; Owens and Koeter 2003; Owens et al. 2003) (Figure 1). The selected uterotropic assays for estrogens and antiestrogens use either the intact juvenile or the castrated ovariectomized adult/juvenile female rat. Both oral and subcutaneous dosing also were considered.

**Hershberger Assay**

The second in vivo assay in T1S, the Hershberger assay, detects antiandrogenic activity simply by weighing androgen-dependent tissues in the castrated male rat (Gray 1998b; Gray et al. 1997a, 1998b; Hershberger et al. 1953) (Figure 2). In this assay, weights of the ventral prostate, Cowper’s glands, seminal vesicle (with coagulating glands and fluids), glans penis, and levator ani/bulbocavernosus muscles are measured in castrated, testosterone-treated (or untreated) male rats after 10 days of oral treatment with the test compound. This assay is very sensitive for detection of androgens and antiandrogens. Other useful endpoints that help reveal the mechanism of action and specificity of the response include weights of the adrenal, liver, and kidney, and measurements of serum (collected by cardiac puncture) levels of testosterone and luteinizing hormone.

The Hershberger assay has been shown to be much more sensitive and specific to AR-mediated alterations than the assessment of endocrine activity in the intact adult male rat, which does not consistently detect the antiandrogenic activity of several weakly antiandrogenic pesticides like p,p’ DDE and linuron (Cook et al. 1993; O’Connor et al. 1999, 2002). In contrast, these toxicants are easily detected in the Hershberger assay (Lambright et al. 2000; Parks et al. 2000; Yamasaki et al. 2003). Chemicals like finasteride, which inhibit 5 alpha reductase activity, can dramatically reduce male accessory sex gland weight with less effect on the levator ani/bulbocavernosus muscle, which has low levels of this enzyme. Although changes in androgen-dependent organ weights in the Hershberger assay are not necessarily considered to be adverse, we have found that chemicals that are positive in the Hershberger assay often produce adverse effects during puberty and after in utero exposure.

Several phases of the OECD-led effort to standardize and validate the Hershberger assay have been completed. To date, the following phases have been completed: Phase 1a (an interlaboratory study using subcutaneous testosterone propionate (17 laboratories, 6 dosage levels); Phase 1b (using flutamide in 6 laboratories at 6 dosage levels); and Phase 2 (17 laboratories, 7 chemicals at several dosage levels, with 3-8 laboratories per chemical) (Yamasaki et al. 2003). The results of these efforts are being prepared for publication.

Both the rat uterotropic and Hershberger assays have been used for several decades to screen chemicals for estrogenicity and androgenicity (Dorfman 1962). Because es-

---

**Figure 1** The Organisation for Economic Cooperation and Development (OECD) Uterotropic Assay Protocol. OECD has standardized the following four uterotropic assay protocols for detection of estrogens and antiestrogens. Protocol A: Intact immature rats were dosed orally for 3 days. Protocol B: Intact immature rats were dosed subcutaneously for 3 days. Protocol C: Ovariectomized young adult rats were dosed subcutaneously for 3 days. Protocol D: Ovariectomized young adult rats were dosed subcutaneously for 7 days. N = 6 female rats per group.

**Figure 2** The Organisation for Economic Cooperation and Development (OECD) Hershberger Assay Protocol is currently being evaluated in a series in interlaboratory studies to detect androgens and antiandrogens. N = 6 castrate-immature male rats per group.
trogens are required for uterine growth and androgens for sex accessory gland growth in both rats and humans, chemicals that are positive in the rat assays for these endocrine activities can be expected to produce predictable responses in humans. These endocrine pathways are highly conserved; in fact, many drugs with antiestrogenic activity or antiandrogenic activity have been shown to produce the anticipated responses in both species. Chemicals examined recently by OECD in these two assays are listed in Table 2.

**Pubertal Female Rat Assay**

The third in vivo mammalian/rat assay included in the proposed EDSTAC T1S battery, the pubertal female rat assay, has been used in our laboratory for nearly two decades (Gray et al. 1988a,b, 1989) (Figure 3). In this assay, weanling female rats are dosed daily by gavage for 21 days while the age at vaginal opening (puberty) is monitored. The females are necropsied at about 42 days of age (reviewed by Goldman et al. 2000). Measurements include serum thyroid hormones, uterine and ovarian weight, and histology. This assay detects alterations in thyroid hormone status, HPG function, inhibition of steroidogenesis, estrogens, and anti-

![Figure 3](image) The Environmental Protection Agency/Endocrine Disruptor Screening and Testing Advisory Committee Pubertal Female Rat Protocol is designed to identify antiestrogenic and antithyroid activities, alterations of hypothalamic-pituitary maturation as related to the onset of puberty, and inhibition of steroidogenesis. N = 15 weanling female rats per group.

### Table 2 Chemicals examined in protocols of the Organisation for Economic Cooperation and Development (OECD) and the US Environmental Protection Agency (EPA), using standardized assays with oral or subcutaneous dosing

<table>
<thead>
<tr>
<th>OECD uterotrophic assay</th>
<th>(four protocols—all successfully identified each chemical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethinyl estradiol: Positive estrogenic control</td>
<td></td>
</tr>
<tr>
<td>Genistein: Phytoestrogen</td>
<td></td>
</tr>
<tr>
<td>Methoxychlor: Estrogenic and antiandrogenic pesticide</td>
<td></td>
</tr>
<tr>
<td>Nonylphenol: Estrogenic “inert” ingredient and surfactant</td>
<td></td>
</tr>
<tr>
<td>o,p’ DDT: Estrogenic pesticide</td>
<td></td>
</tr>
<tr>
<td>Bisphenol A: Estrogenic plastic monomer</td>
<td></td>
</tr>
<tr>
<td>ZM 189,154: Antiestrogenic pharmaceutical</td>
<td></td>
</tr>
<tr>
<td>Dibutyl phthalate: Negative control; plasticizer that inhibits testis Leydig cells</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OECD Hershberger assay</th>
<th>(each chemical successfully identified)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testosterone propionate: Androgenic positive control</td>
<td></td>
</tr>
<tr>
<td>Trenbolone: Androgenic veterinary pharmaceutical</td>
<td></td>
</tr>
<tr>
<td>Methyldihydrotestosterone: Potent androgenic pharmaceutical</td>
<td></td>
</tr>
<tr>
<td>Linuron: Antiandrogenic herbicide</td>
<td></td>
</tr>
<tr>
<td>p,p’ DDE: Antiandrogenic pesticide metabolite</td>
<td></td>
</tr>
<tr>
<td>Flutamide: Potent antiandrogenic pharmaceutical</td>
<td></td>
</tr>
<tr>
<td>Finasteride: Potent inhibitor of enzyme 5 alpha reductase required for DHT synthesis and full androgen-dependent growth of some sex accessory tissues</td>
<td></td>
</tr>
<tr>
<td>Vinclozolin: Antiandrogenic fungicide</td>
<td></td>
</tr>
<tr>
<td>Procymidone: Antiandrogenic fungicide</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EPA in utero-lactational protocol</th>
<th>(partially successful in identifying estrogenicity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methoxychlor: Estrogenic and antiandrogenic pesticide</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EPA execution of adult intact male assay</th>
<th>(did not successfully identify endocrine activity of either chemical)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linuron: Antiandrogenic herbicide</td>
<td></td>
</tr>
<tr>
<td>Methoxychlor: Estrogenic and antiandrogenic pesticide</td>
<td></td>
</tr>
</tbody>
</table>

---

*b* Yamasaki et al. 2003 (see text).
found to be highly reproducible and very sensitive to certain endocrine activities including estrogenicity, inhibition of steroidogenesis, and antithyroid activity (Table 3).

**Alternative Screening Assays**

Alternative in vivo assays also were discussed by EDSTAC and are being evaluated by the EPA. If they are of sufficient sensitivity, specificity, and relevance, they might replace or augment current T1S assays. However, whether they meet such criteria remains to be determined.

**Pubertal Male Rat Assay**

One promising alternative assay is the pubertal male rat assay (Stoker et al. 2000), which detects alterations of thyroid function, HPG maturation, steroidogenesis, and altered steroid hormone function (androstenone). Intact weanling males are exposed to the test substance for approximately 30 days. The age at puberty is determined by measuring the age at preputial separation, and reproductive tissues are evaluated and serum taken for optional hormonal analyses (Figure 4).

The studies conducted on contract for EPA using the pubertal male assay also are presented on the EPA EDSP web site. This assay produced reproducible responses among different laboratories and was sensitive to androgenic effects. A summary of the key chemicals evaluated is presented in Table 3.

**Figure 4** The Environmental Protection Agency/Endocrine Disruptor Screening and Testing Advisory Committee Pubertal Male Rat Protocol is designed to identify antiandrogenic and antithyroid activities, alterations of hypothalamic-pituitary maturation as related to the onset of puberty, and inhibition of steroidogenesis. N = 15 weanling male rats per group.

The studies conducted on contract for EPA using the pubertal male assay also are presented on the EPA EDSP web site. This assay produced reproducible responses among different laboratories and was sensitive to androgenic effects. A summary of the key chemicals evaluated is presented in Table 3.

**Table 3 Chemicals evaluated by the US Environmental Protection Agency (EPA) using the standardized pubertal female rat assay in two contract laboratories or key studies performed in other laboratories**

<table>
<thead>
<tr>
<th>Studies performed in Laboratory 1 for the EPA</th>
<th>Studies performed in Laboratory 2 for the EPA</th>
<th>Key published studies performed in other laboratories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethinyl estradiol: Positive estrogenic control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamoxifen: Mixed estrogen agonist-antagonist; successfully identified mixed action</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ketoconazole: Inhibits steroidogenesis; caused ovarian histological changes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methoxychlor: Estrogenic pesticide; successfully detected by accelerated puberty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenobarbital: Alters liver, hypothalamic, pituitary, and ovarian function; delays puberty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methoxychlor: Estrogenic pesticide; successfully detected by accelerated puberty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ketoconazole: Inhibits steroidogenesis; caused ovarian histological changes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bisphenol A: Weakly estrogenic plastic monomer; negative for endocrine effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propylthiouracil: Antithyroid agent; lowered T4, increased thyroid-stimulating hormone, and caused thyroid histological changes at low doses that slightly delayed rat puberty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fenamidol: Fungicide that weakly inhibits aromatase; slight delay in puberty, but lowered T4 and retarded growth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atarazine: Herbicide that alters hypothalamic-pituitary function; delays puberty and growth</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

*Pubertal female assay data from contract laboratories and summary slides prepared by EPA scientists are available at [http://www.epa.gov/scipoly/oscpendo/assayvalidation/meetings.htm](http://www.epa.gov/scipoly/oscpendo/assayvalidation/meetings.htm).*

*See text for references.*
gens, antiandrogens, inhibitors of steroidogenesis, and antithyroid activity. The chemicals studied to date in this assay are listed in Table 4.

**In Utero-Lactational Assay**

The EDSTAC recommended that the EPA develop and evaluate an in utero-lactational assay due to the unique sensitivity of the fetal reproductive system to disruption by some toxicants. For example, 2,3,7,8 TCDD (dioxin) alters sexual differentiation of male and female rats and hamsters at dosage levels approximately two orders of magnitude below those required to produce adverse effects in pubertal or adult rats (Gray and Ostby 1995; Gray et al. 1997b,c; Wolf et al. 1999). One version of the proposed in utero-lactational assay now being evaluated by the EPA takes about 80 days and uses approximately 10 litters per group (120-150 pups). In this protocol, androgens and antiandrogens can be detected in approximately 2 to 3 wk, and EDCs with antithyroid activity can be detected in infant or weanling offspring after 4 to 5 wk of maternal treatment.

However, the sensitivity and specificity of this assay to some EDCs are of concern. For example, it appears that the in utero-lactational assay is actually less, rather than more, sensitive to some androgenic compounds than the Hershberger assay (Gray et al. 2001; Wilson et al. 2002; Wolf et al. 2002). In addition, inhibitors of steroidogenesis disrupt pregnancy, a nonspecific effect that may be induced by almost any chemical administered at a maternally toxic dose level (Gray et al. 1999c). Due to the large numbers of animals, long study duration, lack of specificity, and occasional insensitivity, this assay may not be ideal for T1 screening. However, as a testing protocol (T2), it offers several advantages over the standard rat multigenerational test: shorter duration, fewer rats on study, more rats could be evaluated to enhance the study power, and inclusion of more endocrine sensitive endpoints. Hence, the in utero-lactational assay could be used as a T2 or a bridge (T1.5) between T1 and T2.

**Multigenerational Testing: Phase 2—T2T**

In a tiered screening and testing approach, only chemicals that display positive reproducible responses in T1S or T1.5 would be evaluated further in full life cycle or multigenerational tests. In T2 testing, not T1S, issues of dose-response, relevance of the route of exposure, sensitive life stages, and adversity are resolved. For some endocrine activities, the number of sensitive endpoints and F1 offspring examined in these assays should be expanded on a case-by-case basis. Anogenital distance at birth and nipple/areola retention in infant female and male rats should be included in testing of

![Table 4](http://www.epa.gov/scipoly/oscpendo/assayvalidation/meetings.htm)

#### Table 4 Chemicals evaluated by the US Environmental Protection Agency (EPA) using the standardized pubertal male rat assay in two contract laboratories or key studies performed in other laboratories

| Studies performed in Laboratory 1 for the EPA | Flutamide: Potent antiandrogenic drug; delayed puberty, among other effects |
|                                               | Ketoconazole: Inhibits steroidogenesis |
|                                               | Methyltestosterone: Androgenic drug; accelerated puberty |
|                                               | Phenobarbital: Alters liver, hypothalamic, pituitary, and testis function; delayed puberty, among other effects |
|                                               | Vinclozolin: Antiandrogenic fungicide; delayed pubertal puberty, among other effects |
|                                               | Dibutyl Phthalate: Plasticizer that inhibits Leydig cell testosterone function; delayed puberty, among other effects |
| Studies performed in Laboratory 2 for the EPA | Vinclozolin: Antiandrogenic fungicide; delayed pubertal puberty, among other effects |
|                                               | Linuron: Antiandrogenic herbicide; delayed puberty, among other effects |
|                                               | p,p′ DDE: Antiandrogenic pesticide metabolite; delayed puberty, among other effects |
|                                               | Phenobarbital: Alters liver, hypothalamic, pituitary, and testis function; delayed puberty, among other effects |
|                                               | Methoxychlor: Estrogenic and antiandrogenic pesticide; reduced androgen-dependent tissues weights |
|                                               | Ketoconazole: Inhibits steroidogenesis |
|                                               | Atrazine: Herbicide, alters hypothalamic pituitary function; delays puberty and growth |

Key published studies performed in other laboratories

- Vinclozolin: Androgen receptor antagonistic fungicide; delayed puberty, increase serum T and luteinizing hormone (Monosson et al. 1999)\(^a\)
- Cyproterone acetate: Antiandrogenic drug; delayed puberty, among other effects
- Polychlorinated dibenyl ether, DE71; antithyroid toxicant; affected thyroid endpoints and delayed puberty (Stoker et al. 2004a,b)\(^b\)
- Finasteride: Potent inhibitor of enzyme 5 alpha reductase required for dihydrotestosterone synthesis; reduced sex gland weight (Marty et al. 2001)\(^b\)

\(^a\)Pubertal female and pubertal male assay data from contract laboratories and summary slides prepared by EPA scientists are available at http://www.epa.gov/scipoly/oscpendo/assayvalidation/meetings.htm.

\(^b\)See text for references.
androgens and antiandrogens, respectively, because they are sensitive, potentially permanent effects that are highly correlated with malformations and reproductive organ weight changes later in life (McIntyre et al. 2001). These early alterations constitute part of the antiandrogen-induced developmental syndromes. The syndrome induced by the AR antagonists differs to some degree from that induced by the phthalates, which inhibit the synthesis of fetal testosterone and insulin-like 3 peptide (insl-3) hormones.

Developmental Syndromes and Tailored Testing

A careful evaluation of the male rat offspring allows one to distinguish the “phthalate syndrome,” in which effects on reproductive development involve a decrease in fetal testicular testosterone and insl-3 peptide hormone biosynthesis (Gray et al. 2002; Mylchreest and Foster 2000; Mylchreest et al. 1999; Parks et al. 2001; Wilson et al. 2004) from the “AR antagonist syndrome,” which is typified by vinclozolin (Gray et al. 1994, 1999b; Lim et al. 2001) or flutamide (McIntyre et al. 2001). The main distinction is that the phthalate syndrome includes testicular, epididymal, and gubernaculal cord agenesis. These lesions are rarely seen in the AR antagonist syndrome, even when all of the males display hypospadias. In addition, of all chemicals that interfere with the androgen signaling pathway in the fetal male rat, only the phthalates affect Leydig cell insl-3 hormone synthesis and cause undescended testes due to gubernaculal agenesis (Wine et al. 1997).

In contrast to the antiandrogens, which primarily affect the male offspring, in utero exposure to androgens has more severe consequences for the female offspring. It is important to consider this information when tailoring T2T based on T1S results. For example, when the female rat fetus is exposed to testosterone or the veterinary pharmaceutical trenbolone in utero, agenesis of the vagina and nipples also is seen at a low rate in the lower dosage groups. In fact, most of the low-dose effects of androgens in the female offspring (retained prostatic and vesicular tissues, nipple agenesis) are effects that are likely to be missed in a standard necropsy.

When conducting a multigenerational study, it is critical to evaluate all of the components of a syndrome so that affected animals are not misidentified and classified as “normal.” In addition, it is important to summarize the data in a manner that clearly delineates the proportion of animals that are affected, indicating that they display any lesion (histological or gross pathology) consistent with the syndrome. In teratology studies, data are typically presented and analyzed in this manner, indicating the number of malformed/number observed on an individual and litter basis, whereas multigenerational studies are frequently presented and analyzed differently, even when clear teratogenic and other developmental responses are noted after birth. Multigenerational protocols are used in T2 because only these protocols expose the animals during all critical stages of development and examine reproductive function of offspring after they mature.

Statistical Limitations of Current Tests

Although the new EPA multigenerational test provides for a comprehensive evaluation of the P0 or parental generation, too few F1 animals (offspring with developmental exposure) are examined after maturity to detect anything but the most profound reproductive teratogens (Gray and Foster 2004) (information regarding the new EPA test is available at http://www.epa.gov/opptsfors/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Drafts/).

P0 animals within a dose group typically respond in a similar fashion to the chemical exposure; however, the response to toxicants in utero can vary greatly even within a litter with only a few animals displaying severe reproductive malformations in the lower dosage groups. For example, dose-related adverse reproductive effects are seen in less than 10% of the F1 offspring with the phthalate DINP (male reproductive malformations, Gray et al. 2000); 2,3,7,8 TCDD (permanent vaginal threads and ovarian atrophy, Gray et al. 1997c); in utero busulfan (ovarian atrophy and delayed puberty, Gray and Ostby 1998); DEHP (epididymal and testicular lesions, and linuron, McIntyre et al. 2000); epididymal and testicular hypoplasia, and androgens (vaginal and nipple agenesis, Wolf et al. 2002). Thus, a standard multigenerational protocol that examines one F1 animal per sex, per litter after maturity from 20 litters per dosage group can detect statistically significant alterations only when they are displayed by 25% or more of the offspring. Histopathological alterations must be displayed by 50% or more of the offspring because a histological examination of the reproductive tract is required only in 10 F1 animals per sex, per dose—far too few to detect anything but the most profound effects on reproductive development. In this regard, there is uncertainty associated with lowest observed adverse effect levels (“LOAELs”) and no observed adverse effect levels (“NOAELs”) identified in multigenerational studies of chemicals that do not include either a complete assessment of all of the endpoints that constitute the syndrome or those included in the EPA new Multigenerational Test Guidelines (http://www.epa.gov/opptsfors/OPPTS_Harmonized/870_Health_Effects_Test_Guidelines/Drafts/).

In our “transgenerational” protocols (Figure 5) (Gray 1998a; Gray and Ostby 1995; Gray et al. 1994, 1997c, 1999b,c, 2002; McIntyre et al. 2000, 2001; Mylchreest et al. 1998, 2000; Wine et al. 1997), we typically use fewer litters (7-10 per dose group), but examine all of the animals in each litter. These protocols actually use fewer animals, but provide enhanced statistical power to detect reproductive effects in the F1 generation. Additional factors, besides detection of adverse effects at necropsy or during data analysis and interpretation, limit interpretation of data from the standard multigenerational reproduction test. The life-long ex-
Exposure of both males and females in the F1 generation, which allows one to detect effects induced in utero, during lactation, or from direct exposure after puberty, can found the identification of when the effect was induced (i.e., during adulthood vs. development) or even the affected sex. In our studies, the dosing period normally is terminated near birth or at weaning, which precludes misinterpretation of the developmental origin of reproductive effects. Nevertheless, it is clear that our transgenerational protocols would not be appropriate for EDCs that induce low-dose alterations in the pubertal or adult animal in the P0 generation.

Test Design and Numbers of F1 Animals

It is important to reiterate that the endpoints described above, which are sensitive to antiandrogens or androgens in utero (listed in detail in Gray and Foster 2004), are not sensitive to xenoestrogens or some other EDCs. We are not suggesting that all EDCs should be tested similarly to the androgens or antiandrogens. Testing should be tailored based on the pharmacological activity demonstrated in T1S. In addition, the developing fetus is not always the most sensitive life stage. Some EDCs disrupt pregnancy by altering maternal ovarian hormone production in P0 dams at dosage levels that appear to be without direct effect on the offspring (Gray et al. 1999b). In such cases, we recommend using the standard EPA multigenerational protocol with minor enhancements, or a transgenerational protocol with exposure continued after weaning. We believe that the

Figure 5 Example of the General Transgenerational (or Transgen- erational or In Utero-lactational) Protocol, which the Environmen- tal Protection Agency (EPA) is considering as part of the Endocrine Disruptor Screening and Testing Program. Pregnant P0 dams are treated with the chemical after implantation, through lactation. At weaning, half of the pups in each litter may be continued on the chemical if the intention is to assess the direct effects of the chemical on pubertal development and reproductive function. In the in utero-lactational screening protocol (option not shown in Figure 5), which the EPA is currently evaluating on a contract basis, F1 pup are necropsied after weaning as young adults. When used for testing, F1 pups are necropsied when they are fully mature (shown in Figure 5). F1 males and females can be mated to produce F2, if warranted (option not shown in Figure 5). Several laboratories have used the protocol shown in this figure to assess the developmental effects of endocrine-disrupting chemicals. Sample sizes range from 5 to 10 litters per dose group, but every F1 pup of the affected sex is examined thoroughly after puberty. Hence, the sample size is smaller than a standard multigenerational test guideline study, but the number of F1 examined for gross histological and morphological abnormalities and organ weights exceeds the standard test.

Figure 6 Estimate of animal use in the Environmental Protection Agency/Endocrine Disruptor Screening and Testing Advisory Committee Screening and Testing Program based on the numbers of animals used per hundred chemicals screened. This estimate is based on several assumptions, including that chemicals would be evaluated in all three in vivo screening assays. The Tier 1 Screen (T1S) includes three in vivo assays (uterotropic, N = 6/group; Hershberger, N = 6/group; and pubertal female rat, N = 15/group, or a total of 27 rats per group). We also assumed that each chemical would be evaluated using three groups (control-, medium-, and high-dosage levels = 81 rats); the high-dosage level should not exceed an MTD (less than a 10% body weight reduction versus control) with a few animals being used for limited dose-range finding for T1S. Based on these assumptions, Tier 1 Screening (T1S) would use approximately 100 animals per chemical screened.

We assumed that about 5% of the chemicals screened would be positive because they act in an antiestrogenic, antiandrogenic, antithyroid manner; or they would inhibit steroidogenesis or hypothalamic-pituitary-gonadal maturation. We also assumed that another 5% would be statistical “false-positives.” We propose that false-positives could be virtually eliminated from entering Tier 2 Testing (T2T) and sent to the “hold” box by repeating or expanding positive T1S results before proceeding to the testing phase of the battery. T2T is the most intensive phase of the battery in terms of animal use, using about 1500 animals in a multigenerational test per chemical. Note that for each 5% increase in the rate of false-positives, there will be an increase of approximately 250 rats used per 100 chemicals in the replications phase between T1S and T2T. Hence, if the false-positive rate were 50%, rather than 5%, animal use would increase from 18,000 only by 2250 rats, to 20,250 per 100 chemicals.
transgenerational or in utero-lactational protocols fill a gap in the testing program for EDCs that should be used only on a case-by-case basis, as indicated by the results of T1S and any T1 repeat.

Scientists are attempting to minimize animal use by using as few animals as possible in the most precise and sensitive assays, by incorporating sensitive in vitro assays in T1S, and by using QSAR models or HPTS in the prioritization of chemicals for screening. It is also possible to avoid unwarranted animal use because chemicals negative in T1S are not subject to T2T. In addition, testing statistical “false-positives” can be almost entirely eliminated by ensuring that T1S assay results are replicated in T1.5 before moving to T2T (Figure 6). Attempts to enhance T2T by adding more sensitive endpoints and a more thorough assessment of the animals already on study also may lead to additional reductions in animal use.

Conclusion

In summary, the EPA EDSP is considering several mammalian and lower vertebrate in vivo and in vitro assays for T1 screening. The in vitro assays include (1) mammalian ER and AR receptor binding and/or transcriptional activation, and (2) steroidogenesis (both testosterone and estradiol/aromatase synthesis assays). In vivo assays that are being evaluated in the laboratory include the following: (1) a rat uterotropism assay, (2) a 10-day rat Hershberger assay, (3) a 20-day puberal female rat assay, (4) a 30-day puberal male rat assay, (5) an in utero-lactational rat assay, (6) a short-term reproduction assay using the fathead minnow, and (7) a frog metamorphosis assay. It is important to retain flexibility in the selection of new assays and endpoints for the screening and testing program so that new methods can be used as replacements or to augment the assays if they offer distinct advantages over the current battery of assays. Enhancements to current life cycle and multigenerational tests also are being considered in an effort to improve the quality of the data on EDCs that will be used for risk assessments.

References


